ABSTRACT OF THE DISCLOSURE

A method of comparing nucleic acid sequences being ESTs included in a first database of sequences and nucleic acid sequences included in a second database of sequences to form groups of sequences from the two databases that all relate to the same gene. For each one or more n-groups of sequences of one of the two databases, associating therewith lists of nucleic acid sequences, each from one of said two databases, each sequence on the list containing the n-groups, and matching sequences on the lists to generate said group.